



## UvA-DARE (Digital Academic Repository)

### The exceptional loss of the pronoun T

Aalberse, S.; Stoop, W.

**DOI**

[10.1016/j.pragma.2015.07.003](https://doi.org/10.1016/j.pragma.2015.07.003)

**Publication date**

2015

**Document Version**

Final published version

**Published in**

Journal of Pragmatics

**License**

Article 25fa Dutch Copyright Act (<https://www.openaccess.nl/en/policies/open-access-in-dutch-copyright-law-taverne-amendment>)

[Link to publication](#)

**Citation for published version (APA):**

Aalberse, S., & Stoop, W. (2015). The exceptional loss of the pronoun T. *Journal of Pragmatics*, 88, 190-201. <https://doi.org/10.1016/j.pragma.2015.07.003>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

# The exceptional loss of the pronoun T

Suzanne Aalberse <sup>a,\*</sup>, Wessel Stoop <sup>b</sup>

<sup>a</sup> University of Amsterdam, Netherlands

<sup>b</sup> Radboud University, Netherlands



## Abstract

Calculations on lexical stability show that the V-pronoun is variable but the T-pronoun is stable across language families (cf. Swadesh 1971 [2006]; Wichmann and Holman, 2009; Tadmor et al., 2010). In fact, the T-pronoun is in the top 10 of basic vocabulary items. Nevertheless, both English and Dutch have lost their T-pronoun. We suggest that this unexpected loss of T in English and Dutch can be explained via an exceptional combination of circumstances, namely (1) a focus on negative politeness, (2) the possibility of deflection via the loss of the T-pronoun and (3) pressure on the inflectional system due to language contact involving adult second language learners. These factors lead to two predictions: one related to the subject-non-subject ratio, and one related to text style. We test both predictions in a corpus of 13th and 16th century texts.

© 2015 Elsevier B.V. All rights reserved.

**Keywords:** Address terms; Inflection; Text style; Language contact; Language change

## 1. Introduction

Languages frequently have two (or more) terms of address. One is familiar or intimate and the other is distant or polite (Brown and Gilman, 1960). Following Brown and Gilman (1960) we will refer to the intimate or familiar pronoun as T (derived from Latin *tu*) and to the distant or polite pronoun as V (derived from Latin *vos*).

Table 1 gives an overview of T- and V-pronouns in various European language families.<sup>1</sup> Pronouns that are still used today are indicated in bold, address forms printed in italics are no longer in use.

The fifth column in this table, 'origin of V', deserves some clarification. The label 2p is used for polite singular pronouns that derive from the second person plural. We see that almost all languages listed have, at some point, used a second person plural form as a singular polite form. In some languages, such as French (*vous*), Turkish (*siz*) and Finnish (*te*), the second person plural pronoun is still used as a polite singular form of address. Other languages have replaced the 2p polite form by other polite forms. The labels 3s and 3p refer to the use of a third person singular and third person plural pronoun as polite singular forms of address.

Following Jucker and Taavitsainen (2003:4), pronouns that derive from titles or from 'abstractions of address' are marked with the label 'respectful title'. Examples of respectful titles are *My Lord*, *My Lady*, *Master* and *Mistress*. Examples of

\* Corresponding author. Tel.: +31 20 5254721.

E-mail address: [s.p.aalberse@uva.nl](mailto:s.p.aalberse@uva.nl) (S. Aalberse).

<sup>1</sup> The information in the table for the West-European languages is based primarily on the work of Jucker and Taavitsainen (2003:4). Information on most other East-European languages is based on Berger (1998:312–313). Additional information on address terms is taken from Bentivoglio (2003:179–180) on Spanish, from Mühlhauser and Harré (1990:136–137) on Italian, from De Groot (2009) on Hungarian, from Betsch (2003:141) on Czech, from Braun (1988:43, 52, 100, 305–306) on Polish, Georgian and Rumanian, from Tegnér (1889, 1922) in Braun et al. (1986:212) on Swedish, from Baden (1804) in (Braun et al., 1986:13) on Danish, from Tylden (1958, 1964) in (Braun et al., 1986:216) on Norwegian, from Simon (2003a, 2003b) on German and from Wales (Wales, 1983, 1996) on English.

Table 1  
Address terms in European languages.

Language family	Language	T	V	Origin of V
Romance	Spanish	<b>tu</b>	<i>vos</i> <i>el/ella</i> <b>usted</b>	(2p) 3s masc/fem Respectful title
Romance	Italian	<b>tu</b>	<i>vos</i> <b>Lei</b>	2p Respectful title
Romance	Portuguese	<b>tu</b>	<i>vos</i> <b>voçe</b>	2p Respectful title
Romance	French	<b>tu</b>	<b>vous</b>	2p
South Slavonic	Bulgarian	<b>тие (tie)</b>	<b>вие (vie)</b>	2p
South Slavonic	Macedonian	<b>тие (tie)</b>	<b>вие (vie)</b>	2p
South Slavonic	Serbian	<b>ти (vi)</b>	<b>ви (vi)</b>	2p
South Slavonic	Croatian	<b>ty</b>	<b>vy</b>	2p
South Slavonic	Bosnian	<b>ti</b>	<b>vi</b>	2p
South Slavonic	Slovenian	<b>ti</b>	<b>vi</b> <i>oni</i>	2p 3p
West Slavonic	Czech	<b>ty</b>	<b>vy</b> <i>pan</i> <i>vašnost</i>	2p Respectful title Respectful title
West Slavonic	Slovakian	<b>ty</b>	<b>vy</b> <i>oni</i>	2p 3p
West Slavonic	Lower Sorbian	<b>ty</b>	<b>wy</b>	2p
West Slavonic	Upper Sorbian	<b>ty</b>	<b>wy</b>	2p
West Slavonic	Polish	<b>ty</b>	<b>wy</b> <i>pan</i> <i>(oni)</i>	2p Respectful title 3p (dialectal)
West Slavonic	Polabian	<b>ty</b>	<b>jaj</b>	Middle low German 2p
East Slavonic	Russian	<b>ты (ty)</b>	<b>вы (vy)</b>	2p
East Slavonic	Belarusian	<b>ты</b>	<b>вы (vei)</b>	2p
East Slavonic	Ukrainian	<b>тие</b>	<b>вие (vie)</b>	2p
Uralic	Finnish	<b>sinä</b>	<b>te</b>	2p
Uralic	Hungarian	<b>te</b>	<b>maga</b> <b>ön</b>	'his body' Self (via German)
Turkic	Turkish	<b>sen</b>	<b>siz</b>	2p
West Germanic	German	<b>du</b>	<i>ihr</i> <i>er/sie</i> <b>Sie</b> <i>Dieselben</i>	2p 3s masc/fem sing 3p Anaphoric pronoun
West Germanic	Dutch	<i>du</i> <i>gij/jij</i>	<i>gij</i> <b>u</b>	2p Respectful title
West Germanic	English	<i>thou</i> <i>ye/you</i>	<i>ye/you</i>	2p
North Germanic	Danish	<b>du</b>	<i>i</i> <b>De</b>	2p 3p
North Germanic	Swedish	<b>du</b>	<i>i</i> → <b>ni</b>	2p
North Germanic	Norwegian	<b>du</b>	<b>De</b>	3p

abstractions of address are the use of metonymic forms of address such as *Your Grace*, *Your Honour* and *Your Well-learnedness* (examples taken from Simon, 2006) and metaphoric forms of address such as *Your Highness* and *Your Majesty*.

A striking conclusion that can be drawn from Table 1 is that variation in the V-pronouns is significantly larger than within the T-pronouns. Whereas the origin of the V-pronoun is variable (in several languages there even is evidence for more than one polite pronoun), the T-pronoun was hardly ever replaced, almost all Indo-European languages use a cognate of \**tu* for the T-pronoun (cf. Aalberse, 2004, 2009; Hickey, 2003; Postma, 2011, 2012). This stability is not a coincidence: calculations on word stability show that T-forms are in the top 10 of most stable and most basic forms around the world (cf. Swadesh 1971 [2006]; Wichmann and Holman, 2009; Tadmor et al., 2010).

The only two exceptions in this table are English and Dutch, which lost their pronouns *thou* and *du*, respectively. In this paper, we try to answer why. We will describe existing hypotheses in Section 2, and then introduce our own hypothesis in

Section 3. In Section 4 we describe how we tested this hypothesis by means of a corpus study and in Section 5 we describe the results. Finally, in Sections 6 and 7 we will offer our discussion and conclusions.

## 2. Background: previous hypotheses

### 2.1. Hypothesis 1: the T becomes marked

This first hypothesis regarding T-loss based on existing literature (cf. Quirk, 1974; Wales, 1983, 1996; Braun, 1988; Dik, 1997) builds on the idea that if there are two address forms, one will be neutral and unmarked, while the other is marked. For example, when only a small group of people receive the V-form in a limited number of settings, receiving the T-form is the unmarked situation, whereas receiving the V-form is the marked one.

According to Brown and Gilman (1960) and Burnley (2003), late medieval societies favored negative politeness; they favored showing respect over marking solidarity, and thus late medieval society tended to favor V.<sup>2</sup> This preference for V then caused the group of V-receivers and V-domains to gradually grow: more people expected to receive the form in more situations. The T thus gradually lost its position as a neutral form to the V, a process called *markedness reversal*. T-use started to imply a degree of disrespect for the addressee (Wales, 1983:116, 1996:75; Braun, 1988:59; Dik, 1997:46). In public situations, a T-form was only used if the addressee was clearly inferior to the speaker, or if the speaker was hit by strong emotions such as anger. It is hypothesized that once the T-form is associated with negative emotions (i.e. anger and contempt), that the language loses this pronoun.

### 2.2. Hypothesis 2: the T becomes associated with the lower classes

The second hypothesis regarding T-loss relates to the choice of T as an identity marker. We know the use of polite forms begins in cities (Wales, 1983:117) and with the higher class (Brown and Gilman, 1960:159; Braun, 1988:59). This could cause the T-form to become associated with a rural background and a lack of education and sophistication. Consequently, speakers avoided the use of T in order to avoid undesired connotations – i.e. being associated with a rural area or with the lower classes (cf. Wales, 1983, 1996).

### 2.3. Problems with these hypotheses

Although markedness reversal and identity may help to construct a scenario for the decrease in the use of T, markedness and identity associations cannot explain T-loss completely for all domains, since these factors are not absolute but depend on usage domains. Clyne et al. (2006:312) show for Modern German that there are unmarked T and unmarked V contexts, and “a large ‘grey area’ where the rules for T and V vary according to individual preference”. For example, the use of *du* (T) between students and faculty members is ‘absolut unüblich’ (‘totally uncommon’) and ‘nicht vorstellbar’ (‘unimaginable’), in some departments such as economics and natural sciences, whereas it is common in smaller departments such as archeology (Clyne et al., 2006:299). V is neutral in large departments whereas it might imply distance in small departments.

Walker (2007:46–48) provides an overview of interpretations of markedness values of address pronouns in historical English texts and shows that pronouns that are marked in some registers, such as depositions, are unmarked in work by Shakespeare. Apart from the objection that markedness reversal and identity association are not absolute associations but depend on the context and the register of use, another objection to the idea that these factors would explain T-loss completely is that we would expect T-loss in many other languages, not just in English and in Dutch, but in all languages with major cities described in the table in (1). Our hypothesis is that *du* and *thou* were lost not only because of the socio-pragmatic factors discussed in Sections 2.1 and 2.2 but also because they provided opportunities for the loss of inflection. In the remainder of this paper we will therefore rephrase these two sociolinguistic hypotheses as a condition for T-loss: ‘There need to be sociolinguistic circumstances that strongly reduce the use of T.’ While we consider this to be a necessary condition, we would like to suggest that sociolinguistic circumstances alone are not sufficient to explain complete T-loss, but that two extra conditions also need to be met in order to achieve this. These conditions, which we will discuss momentarily, have to do with (i) whether the loss of T would make the paradigm of verbal inflection more economical and (ii) whether the circumstances created a need for a more inflectional paradigm; these circumstances

<sup>2</sup> One reviewer remarks that we cannot usefully and meaningfully talk about “societies” as a unified consistent entity without reference to factors such geography and socio-economic class. Negative politeness might not have been favored in all classes, but only in the higher literate classes. We agree on this point. However, if not all groups favored negative politeness and the use of referential forms, this makes the question why T was lost even more urgent.

would typically involve language contact situations. In other words, our hypothesis here is that only on the rare occasion that all three conditions are met, T will be lost.

### 3. Hypothesis and predictions

#### 3.1. Our hypothesis

Our hypothesis for the loss of *du* in Dutch and *thou* in English is that both languages, unlike other European languages, met three conditions:

- (a) There need to be sociolinguistic circumstances that strongly reduce the use of T.
- (b) The loss of T should make the paradigm of verbal inflection more economical.
- (c) The circumstances should create a need for a more economical inflectional paradigm; these circumstances would typically involve language contact situations.

The first condition has been described in Section 2, the second condition will be described in Section 3.1.1, the third in Section 3.1.2.

##### 3.1.1. Inflectional economy via T-loss

The second condition for T-loss is that the loss results in a more economical paradigm. How is inflectional economy related to the loss of the T? First of all it is important to note that the loss of the second person pronoun goes hand in hand with the loss of second person marking on the verb. Aalberse (2009) shows that in Middle Dutch, the relation between the pronoun and the verb ending it triggers is very intimate. If there is a pronoun *du* then it combines with a suffix *-s*, and if there is a suffix *-s* it combines with the pronoun *du*. This intimate relation between the pronoun and the verb is also described for English, for example by Kroch (1989:240), who discusses the loss of second person in English as the loss of the combined set (*thou V+st*), and Lass (1999:162), who characterizes the loss of second person singular inflection as a pronoun story: “The fate of the 2s inflection is really part of the pronoun story: it falls away with the you/thou opposition.” We hypothesize that the combined loss of the second person singular pronoun and the verbal affix it combines with is not only a pronoun story, but partly an inflection story.

If a verb ending is under pressure because it is not completely acquired by the speaker, there are two things speakers can do. The first option is to keep using the pronoun and thus risk producing utterances that might be perceived by others to have agreement errors. Alternatively, the speaker can resort to the use of a pronoun with similar meaning that combines with more economical inflection (like the V) and at the same time avoid agreement errors.

Intuitively, the expectation is that speakers would prefer the latter strategy over the former one. This is exactly what is argued in the literature. Both Coveney (2000) and Zilles (2005) describe similar cases in which two pronouns share similar meanings but go with different types of inflection, one pronoun combines with economical inflection and the other with uneconomical inflection. These cases are the competition between *on* and *nous* in French and between *a gente* and *nos* in Brazilian Portuguese, respectively. Both authors find that speakers prefer to replace pronouns rather than make agreement errors. Note that replacing pronouns because they trigger uneconomical inflection is relevant for subject-pronouns only, because they alone trigger inflection.

If the loss of *du* is indeed partly motivated by the desire to avoid uneconomical inflection, the question is how we define economy in inflection. We will follow Aalberse (2009) and Aalberse and Don (2009, 2011), who, based on a number of typological and acquisition studies, claim that not every form of homophony between inflectional affixes in the verbal paradigm is economical. They use the term economical to indicate those syncretisms that decrease the number of contrasts a language learner needs to assume to account for the data. It is generally accepted (Pinker, 1996) that inflectional contrasts are only assumed to be present by first language learners on the basis of overt evidence. Acquisition literature suggests children acquire inflectional contrasts in a specific order (for an overview see also Harley and Ritter, 2002). Economical deflection implies a loss of distinctions that are learned relatively late. The assumption is that since second person is acquired late relative to the acquisition of first and third person, it reduces the acquisition load when second person is identical to the earlier acquired first or third person marking, since it enables the language learner to account for the input data without having to assume the feature addressee. The syncretism between second person singular and second person plural on the other hand is considered uneconomical in those languages that encode plural in third person, because – following the animacy hierarchy – number marking in third person already implies number marking in second person. If second person singular and second person plural are identical, this would not enable the L1 learner to assume fewer inflectional features, thus the syncretism is not economical. Keeping in mind that some syncretisms are economical whereas others are uneconomical, we will now consider how replacement of the traditional second person singular by second person plural will affect the verbal paradigm.

Consider the verbal paradigm for Middle Spanish in (1):

(1a) Middle Spanish (Penny, 1991:151)

	Singular	Plural
1	-o	-mos
2	-as	-áis
3	-at	-an

As the paradigm in (1a) shows, there is a separate verb ending for every possible combination of number and person in Middle Spanish. Replacing the 2s form with the 2p form would result in the following paradigm as shown in (1b):

(1b) Middle Spanish, hypothetical

	Singular	Plural
1	-o	-mos
2	-áis	-áis
3	-at	-an

That is, only number would be neutralized, and only in the context of second person. This same is true for most other languages in Table 1. We show the Middle Scandinavian and Middle French paradigm as examples:

(2a) Middle Scandinavian (Haugen, 1976:302)

	Singular	Plural
1	-a	-um
2	-aR	-iþ
3	-aR	-a

(2b) Middle Scandinavian, hypothetical

	Singular	Plural
1	-a	-um
2	-iþ	-iþ
3	-aR	-a

The Scandinavian examples in (2) show that replacement of T by V would only yield neutralization of person in the context of second person and that it does not lead to person neutralization. The same is true for French as shown in (3b).

(3a) Middle French (Rickard, 1989)

	Singular	Plural
1	-e	-ons
2	-es	-ez
3	-et	-ent

(3b) Middle French, hypothetical

	Singular	Plural
1	-e	-ons
2	-ez	-ez
3	-et	-ent

Unlike the other languages listed in Table 1, German, Dutch and English also show reduction of person features in cases of T-loss, because in these languages the affix V combined with was homophonous with either first or third person as shown in paradigms (4)–(6).

## (4a) Middle High German (Schmidt et al., 2000)

	Singular	Plural
1	-e	-en
2	-est	-et
3	-et	-en

## (4b) Middle High German, hypothetical

	Singular	Plural
1	-e	-en
2	-et	-et
3	-et	-en

## (5a) (Southern) Middle English (Lass, 1999)

	Singular	Plural
1	-∅	-∅
2	-st	-∅
3	-th	-∅

## (5b) (Southern) Middle English, hypothetical

	Singular	Plural
1	-∅	-∅
2	-∅	-∅
3	-th	-∅

## (6a) Middle Dutch (Van Gestel et al., 1992)

	Singular	Plural
1	-e	-en
2	-s	-t
3	-t	-en

## (6b) Middle Dutch, hypothetical

	Singular	Plural
1	-e	-en
2	-t	-t
3	-t	-en

Thus, in most languages listed in Table 1 replacement of second person singular inflection by second person plural inflection yields neutralization of number in the context of second person only. This type of neutralization cannot be explained by acquisition strategies as it does not affect person oppositions, and hence does not lead to a more economical system. In West-Germanic languages, however, replacing the 2s form with the 2p yields another form of neutralization. Because the second person plural affix that V combines with in West-Germanic languages is homophonous with singular first or third person form, replacement of T by V implies the marked second person feature to become identical to the less marked first or third person as shown in paradigms (4)–(6). This type of neutralization where the marked form becomes identical to less

marked forms is in line with acquisition strategies and is therefore economical. The condition that T-loss results in a more economical inflectional paradigm is therefore only met in the West-Germanic languages Dutch, English and German.

### 3.1.2. Circumstances create need for more economical paradigm

The conditions described so far have been met by Dutch and English, but also by German. However, the German T-form *du* is still in use. This is because German did not meet the last condition: there should not only be an option of making the verbal paradigm more economical, the circumstances should also create a need for a more economical paradigm. A need for more economical inflection is associated with the presence of adult second language and second dialect learners. On average, adult language learners are less successful in mastering inflection than young learners (Johnson and Newport, 1991; Hyltenstam and Abrahamsson, 2003; Blom et al., 2006 among many others). This may also have an effect on the acquisition process of the learners of the next generation. The observation that adult learners are less successful in mastering inflection predicts a correlation between the number of adult learners in a language community and deflection. Indeed, we find that the degree of language contact correlates with the degree of deflection (Trudgill, 2011, and references therein).

It is likely that pressure on the inflectional system was not the same in the three languages under discussion (cf. Weerman, 2007; Weerman and Aalberse, 2010) because they were not affected by second language learning and second dialect learning to the same extent.

In order to illustrate that German had less influence from second language and second dialect learners than did English and Dutch let us briefly consider historical information on migration. Both the Southeast of England and the west coast of the Netherlands were pull areas: they attracted migrants. In contrast, Germany was a push area: workers left Germany in order to work in other areas (cf. Lucassen, 1987). The main document that Lucassen (1987) uses to present exact numbers on migration patterns in Europe are taken from a questionnaire reporting on temporary migration in every department of the French empire gathered between 1808 and 1813 (Lucassen, 1987:7–8). Although the questionnaire is from the nineteenth century, the factors that caused to be pull areas such as closeness to the sea, quality of the soil, and the strength of the communication and travel network, were the same as in earlier periods. Moreover, Howell (2006) reports that in 1500 Holland had approximately 45 percent of its population living in cities compared to 3.2 percent in Germany. The insanitary conditions in the cities meant that deaths vastly outnumbered births. The fact that the cities still managed to grow exponentially can only have been due to extremely high levels of immigration (Howell, 2006:208; De Vries, 1984:275). Like cities in the Netherlands, English cities attracted many immigrants. Lass (1999:3) and Wales (1983:118) report large-scale immigration to London and the Southeast starting in the 14th century.

One result of immigration is second language learning and second dialect learning, which in turn put pressure on the inflectional system. If more immigration means more pressure on the system, and pressure leads to deflection, we should see less deflection in German than in Dutch and English. This holds true in the verbal paradigm (English has two suffixes, Dutch has three and German has four) as well as in the nominal domain:

- (a) German has three definite articles (*der, die, das*) Dutch has two (*de* and *het*) and English has one (cf. Weerman, 2007).
- (b) Adjectival inflection is varied in German (choice between  $-\emptyset$ ,  $-e$ ,  $-en$ ,  $-es$ ,  $-er$ ), limited in Dutch (choice between  $-e$  and  $-\emptyset$ ) and absent in English (cf. Weerman, 2001).
- (c) Both English and Dutch have lost their case system, whereas case is (in part) retained in German (cf. Weerman et al., 2013).

We hypothesize that pressure on the inflectional system due to language contact involving adult second language acquisition and second dialect acquisition is a crucial factor to push deflection via pronoun use. Since this pressure was present in Dutch and English (as they were spoken in pull areas) and absent in German (as it was spoken in push areas), we can understand why T was lost in Dutch and English and not in German.

### 3.2. Predictions

We hypothesize that *du* and *thou* were lost not only because of socio-pragmatic factors but also as a means of deflection. Two different sources of change (politeness and deflection) both motivated the same trend: increase of V at the expense of T. The introduction of V as a singular polite pronoun originated in the higher classes and was thus a change from above – a change that speakers apply above the level of consciousness for social gain-, just as in other European countries (cf. Brown and Gilman, 1960; Jucker and Taavitsainen, 2003). However, deflection, like other changes in the underlying grammar, are in general below the level of consciousness and can thus be termed ‘a change from below’ (cf. Labov, 1994).

According to Andersen (2001), changes from above and changes from below are not equally distributed over text-types: changes from above should occur first in formal text types, whereas changes from below occur first in informal text types. V is introduced in Dutch between 1100 and 1300 (Aalberse, 2009). Because this introduction of V is a change from above, a preference for *gi* should be visible in formal text types in the 13th century. In the 16th century the V-form has

become widely used and pressure on the inflectional system has become stronger. We hypothesize that the change from above (introduction of V in the 13th century) transformed into a change from below (deflection 16th century).

The hypothesis that deflection and thus change from below plays a crucial role in the complete loss of T leads to two testable predictions:

1. If avoidance of T is partly motivated by the desire to avoid a difficult inflectional marker in the 16th century, we expect loss of T between the 13th and the 16th century to be strongest in subjects, since only subjects trigger verbal inflection.
2. If in the 13th century the introduction of V was a change from above, we therefore expect the choice for V in the 13th century to occur most in formal text types. T-loss, however, was also motivated by a change from below, namely deflection in the 16th century. We expect increase in V between the 13th and in the 16th century to occur most clearly in informal text-types because of its association with a change from below.

The remainder of this paper will be devoted to testing these predictions.

#### 4. Methodology

The predictions are tested on a corpus of 13th and 16th century texts. All texts used for the corpus were gathered by Jacqueline Evers-Vermeul (Evers-Vermeul, 2005) and Ninke Stukker (Stukker, 2005) from various sources, such as CD-roms (e.g. the CD-roms Middelnedderlands 'Middle Dutch' and Klassieke literatuur 'Classical literature') and the internet (e.g., the project Laurens Jansz. Coster). Details on the corpus can be found in Stukker (2005), Evers-Vermeul (2005) and Aalberse (2009). For each period a corpus of at least 2000 pages of computerized texts was available.

Because earlier studies showed that prose and verse texts often exhibit different linguistic trends, we kept these texts separate, and will treat them separately in the results, discussion and conclusion. We thus end up with four subcorpora: 13th century prose texts, 13th century rhyming verse texts, 16th century prose texts and 16th century rhyming verse texts. For each of the four subcorpora, all T's (the pronoun *du*) and V's (the pronoun *gi*) were extracted manually, in both full and clitic form. Subject or non-subject forms (e.g. vocative, possessive, object) of these pronouns were marked up as such. This is necessary to test the first prediction, regarding the difference between subjects versus non-subjects.

Testing the second prediction required us to operationalize text style. To this end, we first ran a pilot study on modern Dutch, focusing on word length. We created two 5000-word subcorpora from the SoNaR-corpus of modern written Dutch (Oostdijk et al., 2013): one for formal language, based on newsletters, and one for informal language, based on chats. The pilot study showed that mean word length is much shorter in informal texts than in formal ones: 4.6 characters versus 5.36 characters, respectively. Therefore, we took the mean word length of a text as a (rough) indication of whether that text is formal or not, following Warner (2006). In order to implement this assumption in the main corpus, we used a Python script to calculate the mean word length for each text in our main corpus, and then calculated the median word length of all texts. On the basis of the results, the four subcorpora were split into two parts: an informal part (all texts below the median word length) and a formal part (all texts above that median). We thus ended up with eight subcorpora:

	Rhyming verse	Prose
13th century	1. informal 2. formal	3. informal 4. formal
16th century	5. informal 6. formal	7. informal 8. formal

For each of these corpora, the number of address terms (extracted for the previous prediction) were counted.

#### 5. Results

##### 5.1. Prediction 1: subjects versus non-subjects

###### 5.1.1. Prose

		Subject	Non-subject	% Subject
13th century	Du	292	197	59.71
	Gi	114	223	33.82
16th century	Du	19	60	24.05
	Gi	3029	4021	42.96

Here we see that for both *du* and *gi*, the subject-non-subject ratio changes from one time period to the other. *Gi* is mostly used as a non-subject in the 13th century, and this is less so in the 16th century, although this change is relatively small. *Du*, on the other hand, changes drastically; whereas it is mostly used as a subject in the 13th century, it is mostly used as a non-subject (vocative, possessive, object) in the 16th century.

For both address forms, a chi-square test was carried out to see whether the *subject vs non-subject* ratio was influenced by the time period. This turns out to be the case for *du*,  $\chi^2(1, N = 568) = 34.916, p < .001$ , and an effect was also found for *gi*,  $\chi^2(1, N = 7387) = 10.984, p = .001$ . A separate chi-square test, investigating the difference in change between the two time periods, was significant  $\chi^2(1, N = 7123) = 83.841, p < .001$ , showing that the change was considerably larger for *du*.

### 5.1.2. Rhyming verse

		Subject	Non-subject	% Subject
13th century	Du	864	975	46.98
	Gi	3064	4661	39.66
16th century	Du	37	251	12.84
	Gi	2613	4108	38.88

For rhyming verse texts, too, we see that *gi* stays more or less similar in terms of how often it is used as a subject and how often it is used as a non-subject. *Du*, however, changes drastically; whereas it is frequently used as a subject in the 13th century, it is rarely found as a subject in the 16th century.

For both address forms, a chi-square test was carried out to see whether the time period influenced the *subject vs. non-subject* ratio. For *du*, this turns out to be the case  $\chi^2(1, N = 2127) = 118.829, p < .001$ , but not for *gi*,  $\chi^2(1, N = 14,446) = 0.929, p = .335123$ . As with the prose sub corpus, a separate chi-square test was carried out to investigate the difference in change between the two time periods. The result was significant  $\chi^2(1, N = 2655) = 31,526, p < .001$ , showing that the loss of subject address forms was, again, considerably larger for *du*.

## 5.2. Prediction 2: text style

### 5.2.1. Prose

		Du	Gi	% T
13th century	Informal	424	142	74.91
	Formal	59	147	28.64
16th century	Informal	254	5061	5.78
	Formal	12	149	7.45

We see that, although *du* occurs frequently, there is a strong preference for *gi* in formal texts in the 13th century. In informal texts, on the other hand, *du* is extremely frequent. In the 16th century, the relative frequency of *du* has decreased drastically, in particular in informal texts. In fact, for the informal language it has decreased to such an extent that there no longer seems to be a difference between informal and formal text.

For both centuries, a chi-square test was carried out to see whether the *du vs ghi* ratio was influenced by text style. This turned out to be the case for the 13th,  $\chi^2(1, N = 772) = 138.061, p < .001$ , but not for the 16th century,  $\chi^2(1, N = 5476) = 2.419, p = .120$ . This confirms that there is no longer any effect of text style on the occurrence of *du* in 16th century prose texts. A separate chi-square test, investigating the changes between the two time periods, was significant  $\chi^2(1, N = 5138) = 1028.309, p < .001$ , showing that the loss of *du* was not the same for both text styles; that is, the loss of *du* was significantly larger in informal texts.

### 5.2.2. Verse

		Du	Gi	% Du
13th century	Informal	623	2623	19.19
	Formal	1218	3697	24.78
16th century	Informal	243	4498	5.13
	Formal	49	1605	2.96

For rhyming verse texts, we see the opposite pattern: among second person pronouns, *du* occurs frequently in informal text, and even more in formal text in the 13th century. In the 16th century, the relative frequency of *du* has decreased drastically, in particular for formal texts.

For both centuries, a chi-square test was carried out to see whether text style influenced the *du* vs *ghi* ratio. This turned out to be the case for both the 13th,  $\chi^2(1, N = 8161) = 34.948, p < .001$ , and the 16th century,  $\chi^2(1, N = 6395) = 13.165, p < .001$ . A separate chi-square test, investigating the changes between the two time periods, was significant  $\chi^2(1, N = 5516) = 238.206, p < .001$ , showing that the loss of *du* was greater in formal texts than in informal ones.

## 6. Discussion

The results show that our first prediction is borne out for both prose texts and rhyming verse texts: T-subjects are indeed lost at a higher rate than non-subjects. Our second prediction, namely that T loses its association with informal text types, was true for prose texts, but not for rhyming verse texts. We believe the absence of a text type effect in the rhyming texts might be related to the fact that our operationalization of text style does not go together well with rhyming verse texts, where word length is more determined by the meter than by the level of formality. More detailed investigations of text style could perhaps give us a better insight in which texts should be considered formal and informal.

## 7. Conclusion

The loss of T is a rare phenomenon. We suggest that the T-loss in English and in Dutch could be due to an exceptional combination of circumstances, namely (1) a focus on negative politeness, (2) the fact that the replacement of T by V resulted in a more economical paradigm of verbal inflection, and (3) that the circumstances of language and dialect contact created a need for a more inflectional paradigm.

This scenario for T-loss allowed us to construct a number of testable hypotheses, which were confirmed by corpus research. Our first hypothesis was that since only subjects trigger inflection, the T-subject was lost at a higher rate than non-subjects. This prediction was confirmed for both prose texts and rhyming verse texts. Our second prediction related to the relation between text formality and the distribution of T and V. As is generally the case in natural language, the rise of V was associated with change from above. Since changes from above occur in formal text types first, the prediction was that we would find an association between formal text types and the use of V in the 13th century texts. This prediction was confirmed in the prose texts. The association with formal text types was lost by the 16th century. This is in line with the hypothesis that the increase in the use of V in the 16th century was also a means to avoid uneconomical inflection and thus a typical change from below. So in short, the investigation of the association between address term use and text type further supports the hypothesis that T-loss was motivated both by change from above (politeness) as well as by change from below (deflection in a dialect and language contact scenario).

## Acknowledgments

We would like to thank the editors of the volume, Lotte Hogeweg and Helen de Hoop for their valuable comments and helpful suggestions. We would also like to thank Fred Weerman for useful suggestions. Also, we thank two anonymous reviewers for their criticism and help.

## References

- Aalberse, Suzanne, 2004. *Waer bestu bleven?* *Nederlandse Taalkunde* 9, 231–252.
- Aalberse, Suzanne, 2009. *Inflectional Economy and Politeness: Morphology-Internal and Morphology-External Factors in the Loss of Second Person Marking in Dutch*. LOT, Utrecht.
- Aalberse, Suzanne, Don, Jan, 2009. *Syncretism in Dutch dialects*. *Morphology* 19 (1), 3–14.
- Aalberse, Suzanne, Don, Jan, 2011. *Person and number syncretisms in Dutch*. *Morphology* 21 (2), 327–350.
- Andersen, Henning, 2001. *Markedness and the theory of linguistic change*. In: Henning Andersen (Eds.), *Actualization. Linguistic Change in Progress*. John Benjamins, Amsterdam, Philadelphia, pp. 21–57.
- Baden, Jacob, 1804. *Forelaesninger over det danske Sprog eller resonneret dansk Grammatik*. Brummer, Copenhagen.
- Bentivoglio, Paola, 2003. *Spanish forms of address in the sixteenth century*. In: Taavitsainen, I., Jucker, A.H. (Eds.), *Diachronic Perspectives on Address Term Systems. Pragmatics and Beyond New Series*, vol. 107. John Benjamins, Amsterdam, Philadelphia, pp. 178–191.
- Berger, Tilman, 1998. *Äußere Einflüsse und interne Faktoren bei der Herausbildung der slavischen Anredesysteme*. *Die Welt der Slaven* 43, 307–322.
- Betsch, Michael, 2003. *The system of Czech bound address forms until 1700*. In: Taavitsainen, I., Jucker, A.H. (Eds.), *Diachronic Perspectives on Address Term Systems. Pragmatics and Beyond. New Series*, vol. 107. John Benjamins, Amsterdam, Philadelphia, pp. 125–146.
- Blom, Elma, Poliřenská, Daniela, Weerman, Fred, 2006. *Effects of age on the acquisition of agreement inflection*. *Morphology* 16 (2), 313–336.

- Braun, Friederike, 1988. *Terms of Address. Problems of Patterns and Usage in various Languages and Cultures. Contributions to the Sociology of Language*, vol. 50. Mouton de Gruyter, Berlin [etc.].
- Braun, Friederike, Kohz, Armin, Schubert, Klaus, 1986. *Anredeforschung. Kommentierte bibliographie zur Soziolinguistik der Anrede*. Gunter Narr Verlag, Tübingen.
- Brown, Robert, Gilman, Albert, 1960. The pronouns of power and solidarity. In: Sebeok, T. (Ed.), *Style in Language*. MIT Press, Wiley, Cambridge, MA. Reprinted from: Paulston, C.B., Tucker, G.R. (Eds.), *Sociolinguistics. The Essential Readings*, pp. 156–176, pp. 253–276. Reprinted from: Paulston, C.B., Tucker, G.R. (Eds.), *Sociolinguistics. The Essential Readings*, pp. 156–176.
- Burnley, David, 2003. The T/V pronouns in later Middle English literature. In: Taavitsainen, I., Jucker, A.H. (Eds.), *Diachronic Perspectives on Address Term Systems. Pragmatics and Beyond New Series*, vol. 107. John Benjamins, Amsterdam, Philadelphia, pp. 27–45.
- Clyne, Michael, Kretzenbacher, Heinz-Leo, Norrby, Catrin, Schupbach, Doris, 2006. Perceptions of variation and change in German and Swedish address. *J. Sociolinguist.* 10, 287–319.
- Coveney, Aidan, 2000. Vestiges of nous and the 1st person plural verb in informal spoken French. *Lang. Sci.* 22, 447–481.
- De Groot, Casper, 2009. Identifiability and verbal cross-referencing markers in Hungarian. *Linguistics* 47 (4), 997–1019.
- De Vries, Jan, 1984. *European Urbanization 1500–1800*. Harvard University Press, Cambridge, MA.
- Dik, Simon C., 1997. In: Hengeveld, Kees (Ed.), *The Theory of Functional Grammar, Part 1: The Structure of the Clause*. Mouton de Gruyter, Berlin/New York.
- Evers-Vermeul, Jacqueline, 2005. *The Development of Dutch Connectives. Change and Acquisition as Windows on Form-Function Relations. LOT, Utrecht*.
- Harley, Heidi, Ritter, Elizabeth, 2002. Person and number in pronouns: a feature-geometric analysis. *Language* 78 (3), 482–526.
- Haugen, Einar, 1976. *The Scandinavian Languages: An Introduction to Their History (The Great Languages)*. Faber and Faber, London.
- Hickey, Raymond, 2003. Rectifying a standard deficiency. Pronominal distinctions in varieties of English. In: Taavitsainen, I., Jucker, A.H. (Eds.), *Diachronic Perspectives on Address Term Systems. Pragmatics and Beyond, New Series*, vol. 107. John Benjamins, Amsterdam, Philadelphia, pp. 345–374.
- Howell, Robert B., 2006. Immigration and koineisation: the formation of early modern Dutch urban vernaculars. *Trans. Philol. Soc.* 104 (2), 207–227.
- Hyltenstam, Kenneth, Abrahamsson, Niclas, 2003. Maturational constraints in SLA. In: Doughty, C.J., Long, M.H. (Eds.), *The Handbook of Second Language Acquisition*. Blackwell, Oxford, pp. 539–588.
- Johnson, Jacqueline S., Newport, Elissa L., 1991. Critical period effects on universal properties of language: the status of subadjacency in the acquisition of a second language. *Cognition* 39 (3), 215–258.
- Jucker, Andreas H., Taavitsainen, Irma, 2003. Diachronic perspectives on address term systems: introduction. In: Taavitsainen, I., Jucker, A.H. (Eds.), *Diachronic Perspectives on Address Term Systems. Pragmatics and Beyond New Series*, vol. 107. John Benjamins, Amsterdam, Philadelphia, pp. 1–25.
- Kroch, Anthony S., 1989. Reflexes of grammar in patterns of language change. *Lang. Variat. Change* 1 (3), 199–244.
- Labov, William, 1994. *Principles of Linguistic Change*, vol. 1. Internal Factors. Blackwell, Oxford.
- Lass, Roger, 1999. Phonology and morphology. In: Roger Lass (Eds.), *The Cambridge History of the English Language*, vol. III. Cambridge University Press, Cambridge. pp. 56–186 1476–1776.
- Lucassen, Jan, 1987. *Migrant Labour in Europe, 1600–1900 the Drift to the North Sea (D.A. Bloch, Trans.)*. Croom Helm, London, Sydney, Wolfeboro, New Hampshire.
- Mühlhauser, Peter, Harré, Rom, 1990. *Pronouns and People. The Linguistic Construction of Social and Personal Identity*. Basil Blackwell LTD, Oxford.
- Oostdijk, Nelleke, Reynaert, Martin, Hoste, Veronique, Schuurman, Ineke, 2013. The construction of a 500-million-word reference corpus of contemporary written Dutch. In: *Essential Speech and Language Technology for Dutch*. Springer, pp. 219–247.
- Penny, Ralph, 1991. *A History of the Spanish Language*. Cambridge University Press, Cambridge.
- Pinker, Steven, 1996. *Language Learnability and Language Development. With New Commentary by the Author*. Harvard University Press, Cambridge, MA.
- Postma, Gertjan, 2011. Het verval van het pronomen du - dialect geografie en historische syntaxis. *Nederlandse Taalkunde* 16, 56–87.
- Postma, Gertjan, 2012. Clause-typing by the loss of the 2nd person pronoun du 'you' in Dutch, Frisian and Limburgian dialects. In: Camacho-Taboada, V., Jiménez-Fernández, A.L., Martín-González, J., Reyes-Tejedor, M. (Eds.), *Information Structure and Agreement*. Benjamins, Amsterdam, pp. 217–254.
- Quirk, Randolph, 1974. *The Linguist and the English Language*. Edward Arnold, London.
- Rickard, Peter, 1989. *A History of the French Language*. Unwin Hyman, London.
- Schmidt, Wilhelm, Langner, Helmut, Wolf, Norbert R., 2000. *Geschichte der deutschen Sprache. Ein Lehrbuch für das germanistische Studium* 8, Hirzel, Stuttgart [etc.].
- Simon, Horst, 2003a. From pragmatics to grammar. Tracing the developments of respect in the history of the German pronouns of address. In: Taavitsainen, I., Jucker, A.H. (Eds.), *Diachronic Perspectives on Address Term Systems. Pragmatics and Beyond New Series*, vol. 107. John Benjamins, Amsterdam, Philadelphia, pp. 85–123.
- Simon, Horst, 2003b. Für eine grammatische Kategorie >Respekt< im Deutschen. *Synchronie, Diachronie und Typologie der deutschen Anredepronomina (Linguistische Arbeiten)*. Max Niemeyer, Tübingen.
- Simon, Horst, 2006. Reconstructing historical orality in German – what sources can we use? In: Taavitsainen, I., Härmä, J., Korhonen, J. (Eds.), *Dialogic Language Use – Dimensions du dialogisme – Dialogischer Sprachgebrauch. Société Néophilologique (Mémoires de la Société Néophilologique de Helsinki 66)*, Helsinki, pp. 7–26.
- Stukker, Ninke M., 2005. *Causality Marking Across Levels of Language Structure: A Cognitive Semantic Analysis of Causal Verbs and Causal Connectives in Dutch. LOT, Utrecht*.
- Swadesh, Morris, 2006. *The Origin and Diversification of Language Transaction Publishers, New Brunswick, NJ (original work published 1971 – Chicago Aldine Atherton)*.
- Tadmor, Uri, Haspelmath, Martin, Taylor, Bradley, 2010. Borrowability and the notion of basic vocabulary. *Diachronica* 27 (2), 226–246.
- Tegnér, Esaias, 1889. *Tyska inflytelsler på svenskan. Arkiv för nordisk filologi* 5, 155–166, 304–344.

- Tegnér, Esaias, 1922. *Språkets makt över tanken*. In: Tegnér, E. (Ed.), *Ur språkens värld 1: Tre uppsatser*. Bonniers, Stockholm, pp. 165–342.
- Trudgill, Peter, 2011. *Sociolinguistic typology: social determinants of linguistic complexity*. Oxford University Press.
- Tylden, Per, 1958. 2. persons personlege pronomen dualis og pluralis i gamalnorsk og mellomnorsk diplomål. Universitetet i Bergen, Bergen.
- Tylden, Per, 1964. *Fleirtalsform i høfleg tiltale (pluralis reverentiae)*. Det kongelige norske videnskabers selskabs forhandling 37, 33–43.
- Van Gestel, Frank, Twilhaar, Jan Nijen, Rinkel, Tineke, Weerman, Fred, 1992. *Oude zinnen*. Grammaticale analyse van het Nederlands tussen 1200–1700. Martinus Nijhoff, Leiden.
- Wales, Katie, 1983. *Thou and you in early modern English: brown and Gilman re-appraised*. *Stud. Linguist.* 37, 107–125.
- Wales, Katie, 1996. *Personal Pronouns in Present-day English (Studies in English Language)*. Cambridge University Press, Cambridge.
- Walker, Terry, 2007. *Thou and You in Early Modern English Dialogues*. John Benjamins, Amsterdam, Philadelphia.
- Warner, Anthony, 2006. *Variation in the interpretation of change in periphrastic do*. In: van Kemenade, A., Los, B. (Eds.), *The Handbook of the History of English*. Blackwell Publishing, Oxford, pp. 45–67.
- Weerman, Fred, 2001. *Dynamiek in taal en de explosie van de neerlandistiek*. Vossiuspers UvA, Amsterdam.
- Weerman, Fred, 2007. *It's the economy, stupid! Een vergelijkende blik op 'men' en 'man'*. In: Hüning, M., Vogel, U., van der Wouden, T., Verhagen, A. (Eds.), *Nederlands tussen Duits en Engels*. SNL, Leiden, pp. 19–47.
- Weerman, Fred, Aalberse, Suzanne, 2010. *Voorwaarden voor deflexie: verlies van de tweede persoon*. Paper presented at A Germanic Sandwich 2010: Dutch between English and German A comparative linguistic conference, September 18th 2010, University of Oldenburg.
- Weerman, Fred, Olson, Mike, Cloutier, Robert A., 2013. *Synchronic variation and loss of case: formal and informal language in a Dutch corpus of 17th-century Amsterdam texts*. *Diachronica* 30 (3), 353–381.
- Wichmann, Søren, Holman, Eric W., 2009. *Temporal Stability of Linguistic Typological Features*. LINCOP Europa, München.
- Zilles, Ana M.S., 2005. *The development of a new pronoun: the linguistic and social embedding of a gente in Brazilian Portuguese*. *Lang. Variat. Change* 17, 19–53.

### Primary sources

- CD-rom Middelnedderlands 'Middle Dutch' (1998). Sdu, Den Haag/Antwerpen. CD-rom. *Klassieke literatuur: Nederlandse letterkunde van de Middeleeuwen tot en met de Tachtigers* 'Classical literature: Dutch literature from the Middle Ages to the Eightiers' (1999). Het Spectrum, Utrecht. Oostendorp, M. van (Ed.), *Project Laurens Jansz. Coster: Klassieke Nederlandstalige literatuur in elektronische edities* 'Project Laurens Jansz. Coster: Classical Dutch literature in electronic editions'. Available on: <http://cf.hum.uva.nl/dsp/ljc/>. Van der Lem, G.A.C. (Ed.), *De Tachtigjarige Oorlog. Opstand en strijd in de Nederlanden (ca. 1550-ca. 1650)*. Available on: <http://dutchrevolt.leidenuniv.nl/>.